

An end-user oriented data gateway for the european climate observations, modelling and services initiative

ECOMS User Data Gateway

Antonio S. Cofiño

antonio.cofino@unican.es

Santander Meteorology Group
Climate Data Services
Universidad de Cantabria (UC)



Dpto. Matemática Aplicada y
Ciencias de la Computación
Instituto de Física de Cantabria

<http://www.meteo.unican.es>

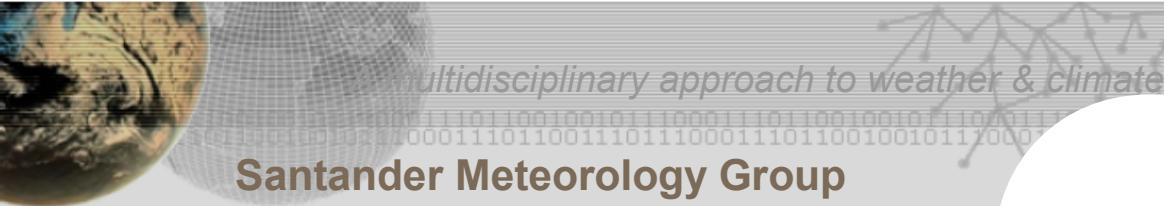


Seasonal-to-decadal climate Prediction for the
improvement of European Climate Services

• EUPORIAS

European Provision Of Regional Impacts
Assessments on Seasonal and Decadal
Timescales

- ECOMS User Data Gateway
- Datasets
- Variables
- How to access data
 - Registration and authorization
 - OpenDAP access
 - R Interface
- Work in Progress



Santander Meteorology Group
A multidisciplinary approach for weather & climate

ECOMS

ECOMS

The European initiative for
climate service
observation and modelling



Seasonal-to-decadal climate Prediction for the
improvement of European Climate Services

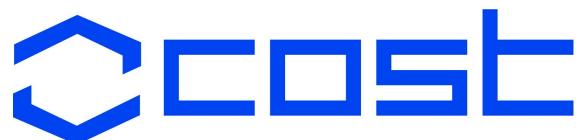


European Provision Of Regional Impacts
Assessments on Seasonal and Decadal
Timescales

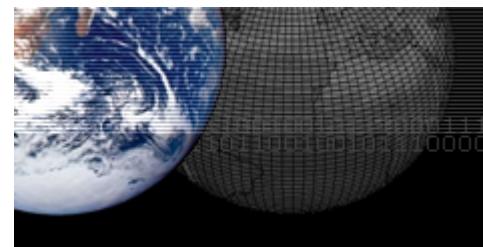
Improve the climate
prediction systems

Improve the usability
and use of prediction

Observations
(decadal scale)



VALUE
COST Action ES1102
Validating and
Integrating
Downscaling Methods
for Climate Change
Research



PUBLIC and
PRIVATE
datasets for
community

CORDEX:
Empirical Statistical
Downscaling.

Datasets of postprocessed predictors statistical downscaling purposes and hi-res observational gridded datasets over different world regions.

Other initiatives

The User Data Gateway (UDG) provides a **homogeneous access end-point** to collections of impact-relevant variables.

The aim of UDG is to gather different data sources with **different terms of use (policies)** in a single data service, so that users can access all the data and metadata they typically need (seasonal forecasts, reanalysis and observations) in a homogeneous and simple way, **without worrying about the inherent complexities of data access, download and post-processing** of the variables stored in massive archive systems at different institutions.

Santander Meteorology Group

A multidisciplinary approach for weather & climate

Wiki doc

The ECOMS User Data Gateway

The European Climate Observations, Modelling and Services initiative ([ECOMS](#)) coordinates the activities of three ongoing European projects ([EUPORIAS](#), [SPECS](#) and [NACLIM](#)). Different activities carried out in these projects require seasonal forecasts from state-of-the-art forecasting systems (e.g. NCEP/CFSv2 or ECMWF/System4) for a reduced number of variables. This information can be obtained directly from the data providers, but the resulting formats, aggregations and vocabularies may not be homogeneous across datasets, thus requiring some post processing. Moreover, different data policies hold for the various datasets —which are freely available only in some cases— and therefore data access may not be straightforward. Thus, obtaining seasonal climate forecast data is typically a time consuming task.

As part of the data management activities in these projects, the *ECOMS User Data Gateway (ECOMS-UDG)* has been developed by the [Santander MetGroup](#) in order to facilitate seasonal forecasting data access to end users. The needed variables have been downloaded from data providers and locally stored in a THREDDS data server implementing fine-grained user authorization (see the available [datasets](#) and [variables](#)). Thus, users can efficiently retrieve the subsets best suited to their particular research aims (for particular regions, periods and/or ensemble members) from a large volume of information. Since [R](#) has been adopted for a number of tasks in these projects (validation, downscaling, etc.), an R package for data exploration and access has been developed ([ecomsUDG.Raccess](#)) and additional functionalities will be developed to meet the users' needs. Moreover, a number of alternative tools are being developed in order to provide different user-friendly interfaces for accessing the information (Python, Matlab, web portals, ...).

This wiki page provides an up-to-date description of the **ECOMS-UDG**, including information on the available datasets and variables, and the documentation of the available tools. The following documents and links are the basic references:

- **Technical notes:**

- [The ECOMS Data Gateway \(v2.0\)](#)
- [Validation of System4 dataset \(v1.0\)](#)

- **Tools:**

- [R package: link to Latest stable release](#)
- [Other interfaces to the ECOMS-UDG](#)

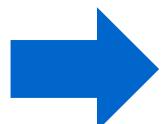
Contents of the wiki page:

1. [THREDDS Data Server \(TDS\)](#)
 - [User registration](#)
 - [Available datasets](#)
 - [Available variables](#)
 - [Exploration via Web](#)
 - [Interfaces for Data Access](#)
 - [DAP access](#)
2. [R Package for Homogenized Data Access](#)
 - [Pre-requisites](#)
 - [Authentication](#)
 - [Data homogenization](#)
 - [Examples](#)

Recent Changes:

- [New stable minor version release of ecomsUDG.Raccess \(v2.2-1\)](#) -- Posted on 2014-08-05 19:14 : author [juaco](#) : categories [Specs](#) [Euporias](#) [Data Server](#) [Ecoms Udg](#)
- [New stable minor version release of ecomsUDG.Raccess \(v2.2-0\)](#) -- Posted on 2014-07-16 13:42 : author [juaco](#) : categories [Specs](#) [Euporias](#) [Data Server](#) [Ecoms Udg](#)
- [New patch release \(v2.1-1\) of ecomsUDG.Raccess package](#) -- Posted on 2014-07-11 16:16 : author [juaco](#) : categories [Specs](#) [Euporias](#) [Data Server](#) [Ecoms Udg](#)
- [ecomsUDG.Raccess v2.1-0 released](#) -- Posted on 2014-07-11 16:14 : author [juaco](#) : categories [Specs](#) [Euporias](#) [Data Server](#) [Ecoms Udg](#)
- [ecomsUDG.Raccess v2.0-0 released](#) -- Posted on 2014-06-16 20:24 : author [juaco](#) : categories [Specs](#) [Euporias](#) [Data Server](#) [Ecoms Udg](#)
- [ecomsUDG.Raccess v1.0-0 released](#) -- Posted on 2014-02-17 13:50 : author [juaco](#) : categories [Specs](#) [Euporias](#) [Data Server](#) [Ecoms Udg](#)
- [Re-branding to ECOMS UDG](#) -- Posted on 2014-01-23 20:06 : author [antonio](#) : categories [Euporias](#) [Data](#) [Specs](#) [Server](#) [Ecoms Udg](#)
- [Defined an initial list of variable](#) -- Posted on 2013-12-13 10:52 : author [gutierrezm](#) : categories [Specs](#) [Euporias](#) [Data](#) [Server](#)
- [The CFSRR dataset has been added to the portal](#) -- Posted on 2013-09-24 12:26 : author [antonio](#) : categories [Euporias](#) [Data](#) [Specs](#) [Server](#)
- [First version of the user data portal](#) -- Posted on 2013-04-25 20:07 : author [antonio](#) : categories [Specs](#) [Euporias](#) [Data](#) [Server](#)

Fully documented.
Feedback welcome.





System4 (provided by ECMWF) [Seasonal forecast]

The [System 4](#) seasonal forecasting system became operational in November 2011. The corresponding hindcast is archived in the Meteorological Archival and Retrieval System ([MARS](#)), the main data repository at the ECMWF, as a collection of GRIB-1 files at 0.75° spatial resolution. The downloaded data has been exposed as three different virtual datasets (see the [available variables](#) for these datasets):

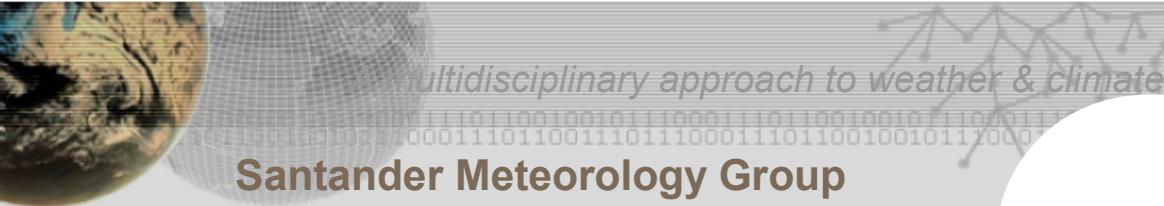
- [System4_seasonal_15](#): There are twelve initializations (hereafter called [runtimes](#)) per year (the first of January, February, ...), each with 15 members running for 7 months (hereafter called simply [times](#)). Period: 1981-2010.
- [System4_seasonal_51](#): There are only four [runtimes](#) per year (the first of February, May, August and November), each with 51 members running for 7 months. Period: 1981-2010.
- [System4_annual_15](#): There are four [runtimes](#) per year each with 15 members, but the forecasts run for 13 months. Period: 1981-2010.

CFSv2 (provided by NCEP) [Seasonal forecast]

The [CFS version 2](#) seasonal forecasting model became operational at NCEP in March 2011. The corresponding [retrospective CFSv2](#) forecast dataset is stored in the [NOMADS server](#) as a collection of GRIB-2 files at 1° spatial resolution. The downloaded data is exposed as a single virtual dataset (see the [available variables](#) for this dataset):

- [CFSv2_seasonal_16](#). There are four initializations (4 cycles) from every 5th day (thus providing on average 24 members per month) running for 9 months (see [CFSv2 members](#) for more detailed information of members' construction for this dataset). Period: 1982-2010. **Note:** For better comparability with other hindcasts, the [R data access package](#) defines by default an ensemble of 16 members for each lead month and forecast season.





Santander Meteorology Group

A multidisciplinary approach for weather & climate

Available datasets

NCEP Reanalysis 1 (provided by NCEP/NCAR) [Reanalysis]

The NCEP/NCAR Reanalysis 1 project provides 6-hourly information of a number of atmospheric variables both at surface and (17) pressure levels. A large subset of this data was downloaded in November 2010 from [PSD](#) in its original format (NetCDF files at 2.5° spatial resolution). The downloaded data is exposed as a single virtual dataset (see the [available variables](#) for this dataset).

- [NCEPReanalysis1](#). Six hourly information of surface and pressure level variables. Period: 1948-2010.

WFDEI (provided by WATCH project) [Observations]

The WFDEI ([WATCH](#) Forcing Data methodology applied to ERA-Interim data) provides 3-hourly and daily averages of eight meteorological variables for the global land surface at 0.5° resolution in NetCDF format. The data was downloaded in February 2014 and it is exposed as a single virtual dataset (see the [available variables](#) for this dataset).

- [WFDEI_daily](#). Daily information of the original and some transformed variables. Period: 1979-2012.

Work in progress...

- 
- [GloSea5](#) (provided by Met Office)
 - [E-OBS](#) (accessed remotely through the OpenDAP server from ECA&D, <http://www.ecad.eu>)
- 

Available variables

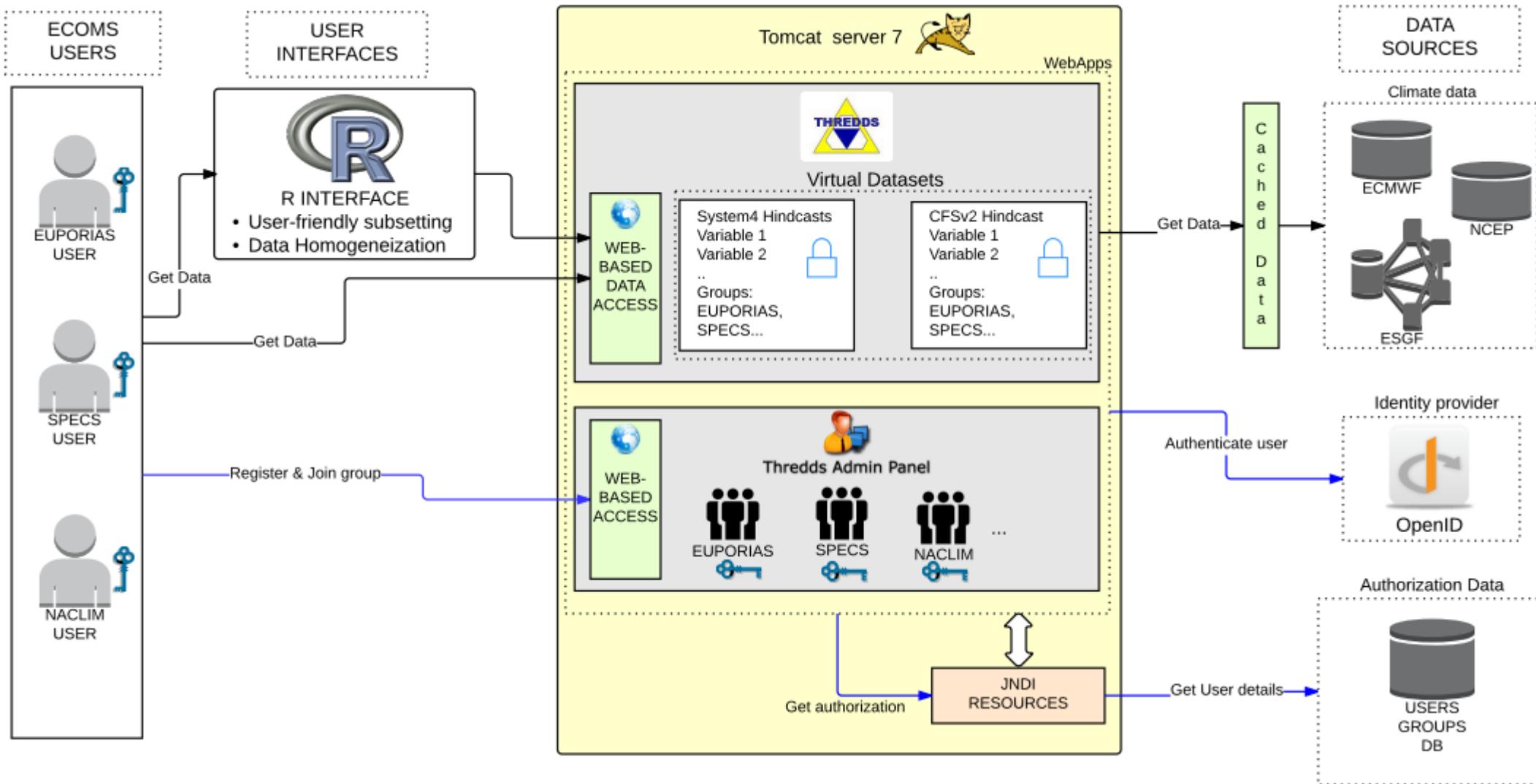
R name	Variable description	Observations:	Reanalysis:	Seasonal forecasting models:				
		WFDEI_daily	NCEP Reanalysis1	System4 seasonal_15	System4 seasonal_51	System4 annual_15	CFSv2 seasonal_16	SPECS-ESGF
<i>Surface variables</i>								
tas	Near-Surface air temperature	DM	6h	6h/DM	DM		e	e
tasmax	Daily Maximum Near-Surface Air Temperature	DX(#)	6h	DX	DX	DX	DX	e
tasmin	Daily Minimum Near-Surface Air Temperature	DN(#)	6h	DN	DN	DN	DN	e
tp	Total precipitation amount	DA	6hA	DAr	DAr	DAr	DA	e
psl	Sea Level Pressure		6h	6h	6h	12h	e	e
ps	Surface air pressure	DM		6h(*)			e	
wss	Wind speed (at 10m)	DM		6h(*)	e	e	e	
tdps	2m Dewpoint Temperature			6h	e			e
huss	Surface (2m) specific humidity	DM	6h	6h(*)			e	



Santander Meteorology Group

A multidisciplinary approach for weather & climate

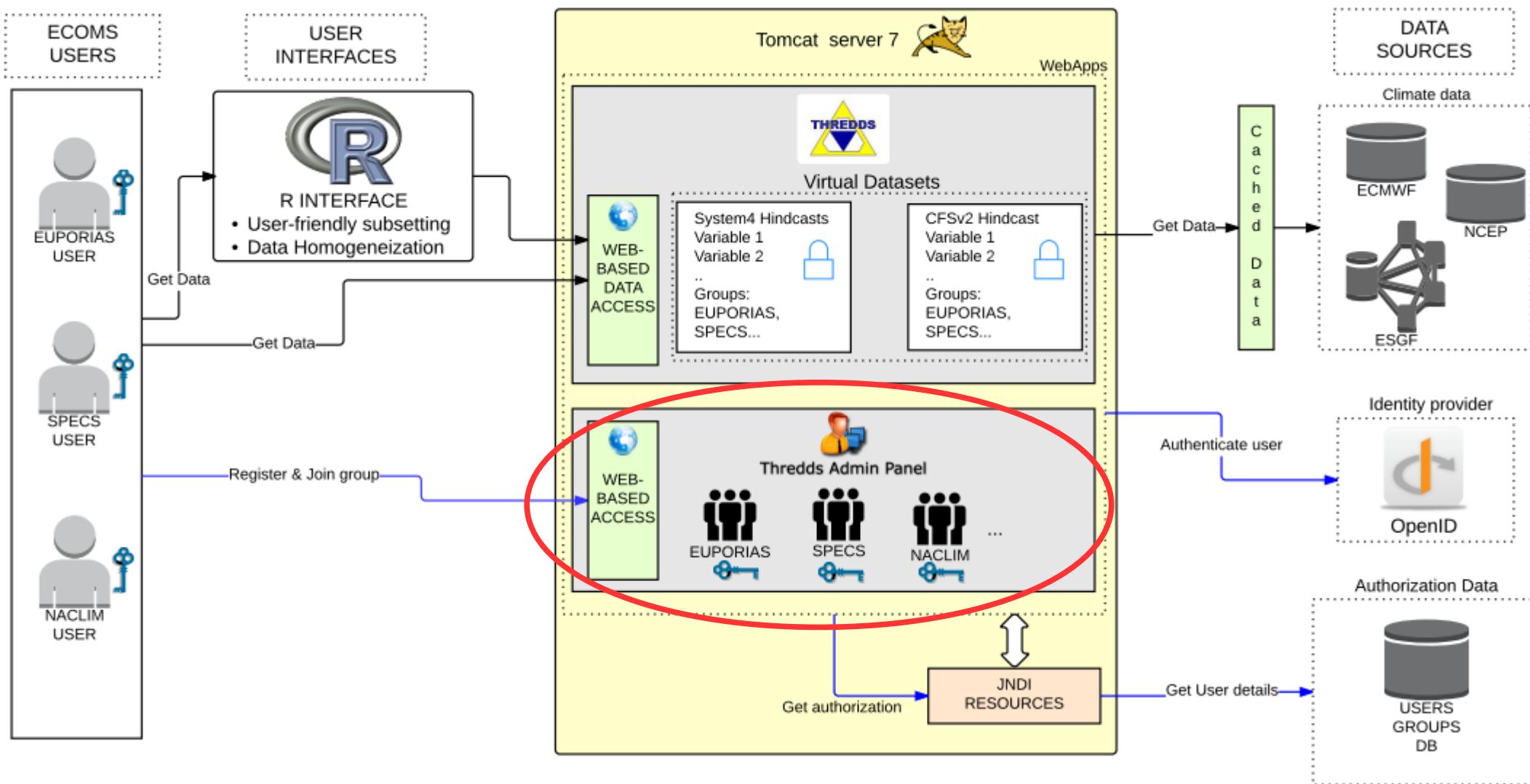
Architecture



Santander Meteorology Group

A multidisciplinary approach for weather & climate

Registration and Authorization



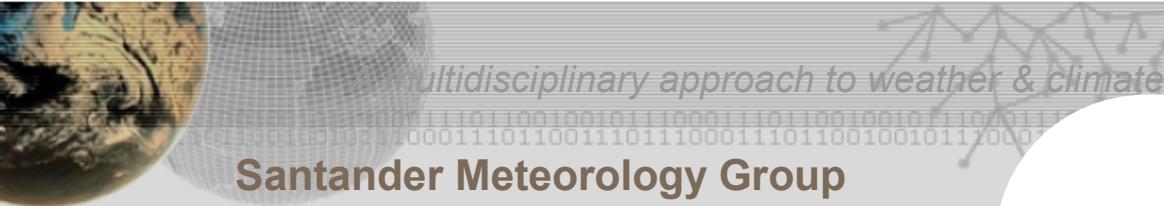
Santander Meteorology Group
A multidisciplinary approach for weather & climate

Registration and Authorization

The registration and authorization process is been managed by the the TAP service

The screenshot shows the Thredds Admin Panel interface. On the left, a sidebar menu includes 'Thredds Admin Panel' (with a user icon), 'Welcome, chus', 'Home', 'Account', 'My groups', 'My Datasets', and 'Support'. The main content area has a title 'My groups' and a 'Group selection' dropdown set to 'SPECS'. Below it is a section titled 'Assigned groups' containing a card for 'EUPORIAS' with details: Name: EUPORIAS, Description: Euporias project, Privacy: Authorization required. Buttons for 'Datasets', 'Policies', and 'Remove' are shown. To the right, a 'My Datasets' section lists two datasets: 'cfsrr' (CFSV2 Reforecast dataset, Open access) and 'system4' (System4 dataset hindcast, Authorization required). Each dataset card includes a 'Policies' button.

The user has to enrol into different projects which are collections of datasets. This enrolment may be supervised by the groups/project administrator.



Policies and terms of use

The enrolment process may also include the acceptance of policies or terms of use imposed by data providers.

A screenshot of a web browser window titled "My groups". The address bar shows "meteo.unican.es/tap/user/groups#". The main content area displays "Datasets and Policies of group PUBLIC_DATA". Under this heading, there is a section for the "Dataset cfsrr" which is described as a "CSV2 Reforecast dataset". Below this, there is a link to a "CFSRR agreement" represented by a document icon. A detailed description of the CFSRR agreement follows, mentioning it is developed at the Environmental Modeling Center at NCEP and includes a restriction. At the bottom, a contact email "cfs@noaa.gov" is provided.

My groups

meteo.unican.es/tap/user/groups#

Thredds Admin Panel

Welcome, antonio

Logout

Datasets and Policies of group PUBLIC_DATA

Dataset cfsrr

CSV2 Reforecast dataset

CFSRR agreement

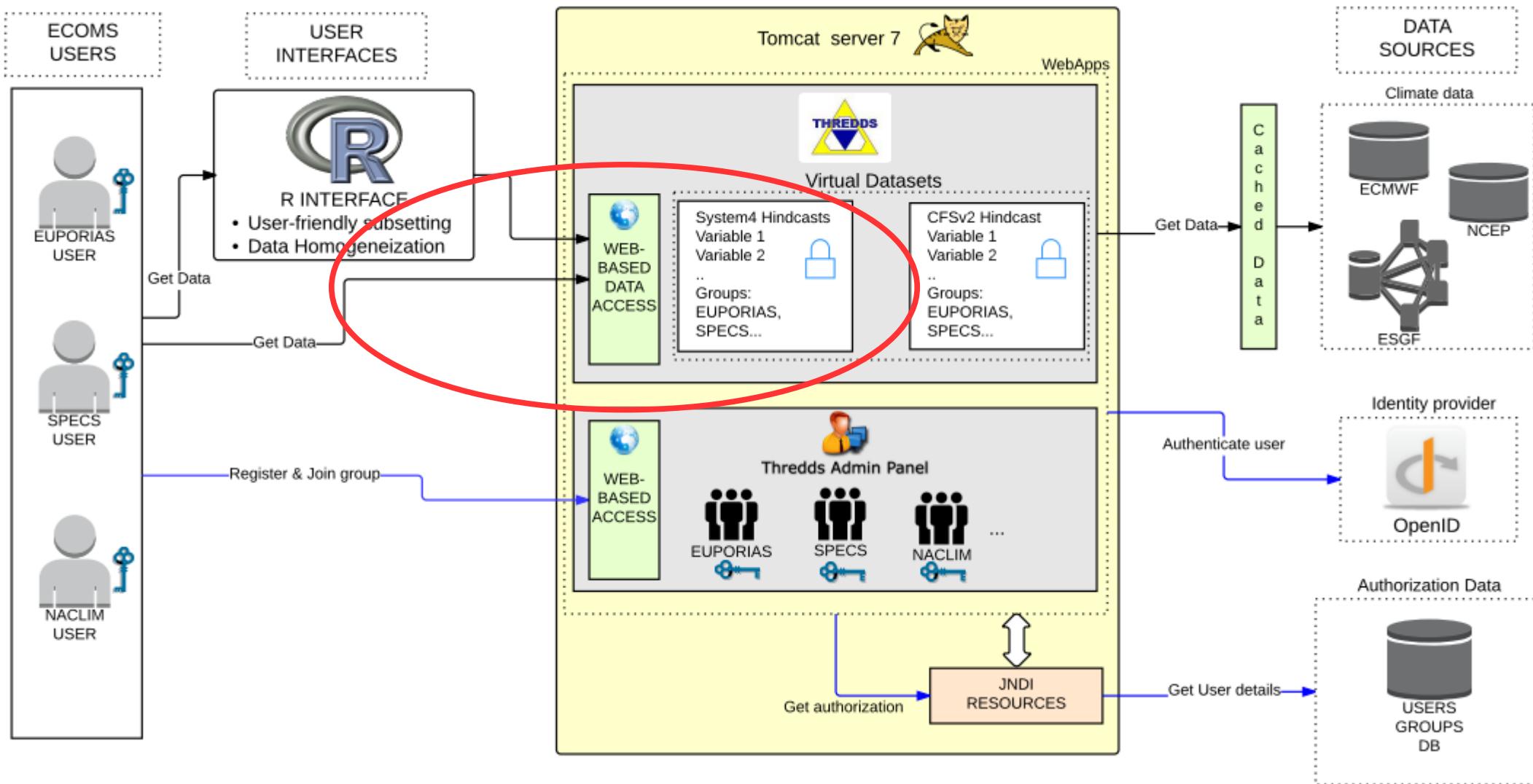
The NCEP Climate Forecast System Version 2 (CFSv2, [http:](http://) developed at the Environmental Modeling Center at NCEP ar restriction.

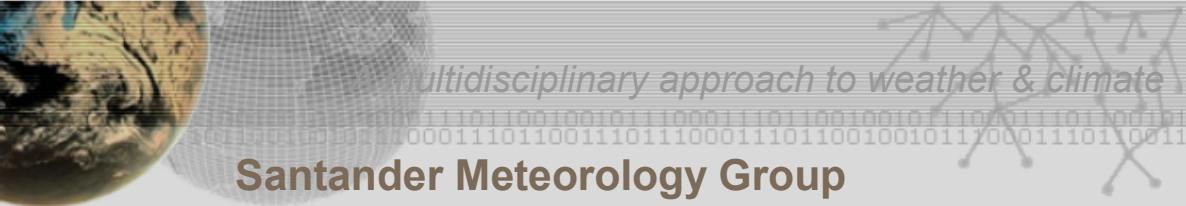
Contact: cfs@noaa.gov

Santander Meteorology Group

A multidisciplinary approach for weather & climate

OpenDAP access





Santander Meteorology Group

A multidisciplinary approach for weather & climate

TdsStaticCatalog... EcomsUdg/DataServ... OPeNDAP Datas... http://w...][0:1:0] User login

www.meteo.unican.es/tds5/catalog.html octave sqlite3

Catalog <http://www.meteo.unican.es/tds5/catalog.html>

Dataset	Size	Last Modified
Reanalysis Datasets	--	
NCEP Reanalysis 1 Datasets/	196.2 Gbytes	2014-06-18T14:08:00Z
ECMWF's Interim Reanalysis Datasets/	61.61 Gbytes	2014-07-09T13:46:09Z
Observations Datasets	--	
WFDEI Datasets/	167.3 Gbytes	2014-06-13T09:43:00Z
EOBS Gridded Datasets/	87.59 Gbytes	2013-06-18T15:30:38Z
Seasonal Forecast Model Datasets	--	
ECMWF's System4 Datasets/	27.58 Tbytes	2014-06-02T12:56:00Z
NCEP's CFSv2 Datasets/	16.49 Tbytes	2014-05-02T12:56:00Z
EC-EARTH V3 Datasets/	1.733 Gbytes	2014-10-01T14:37:00Z

Initial TDS Installation at Santander Meteo Group see Info
THREDDS Data Server [Version 4.3.18 - 20130801.1600] Documentation



Santander Meteorology Group

A multidisciplinary approach for weather & climate

The ECOMS-UDG provides access to datasets with no worries about files. The UDG can offer **views for a dataset and just one endpoint** for the view.

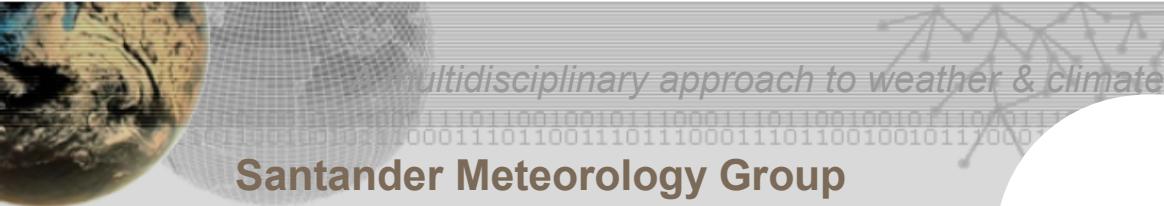
This makes possible to **forget** about **file-name-mangling** and focus on information contained into the dataset.

This is possible thanks to the THREDDS capabilities to **aggregate collections of files as one “NetCDF” entity**.

This approach allows to clients to **materialize** their own **views** to local files. From netcdf-library (>4.1) compiled with opendap support you can do (netcdf-4 file with shuffling and deflate filters):

```
nccopy -k 4 -s -d 6 http://opendap-url.dods?var1[...] local-file_var.nc
```

Init



<http://www.meteo.unican.es/tds5/catalogs/system4/System4Datasets.html>

Authentication Required

A username and password are being requested by <http://www.meteo.unican.es>. The site says:
"EUPORIAS Datasets"

User Name:

Password:

THREDDS Data Server

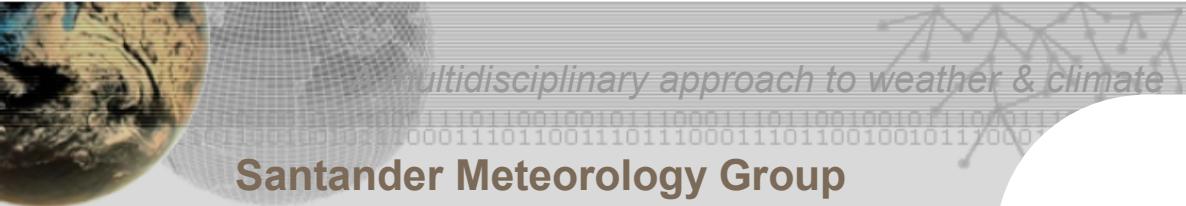
Catalog <http://www.meteo.unican.es/tds5/catalogs/system4/System4Datasets.html>

Dataset: ECMWF's System4 Datasets/System4 15 members Seasonal range Dataset

- *Data format:* GRIB-1
- *Data size:* 23.73 Tbytes
- *Data type:* GRID
- *ID:* system4/System4_Seasonal_15Members.ncml
- *RestrictAccess:* system4

Access:

1. **OPENDAP:** /tds5/dodsC/system4/System4_Seasonal_15Members.ncml
2. **WCS:** /tds5/wcs/system4/System4_Seasonal_15Members.ncml
3. **WMS:** /tds5/wms/system4/System4_Seasonal_15Members.ncml
4. **NetcdfSubset:** /tds5/ncss/grid/system4/System4_Seasonal_15Members.ncml



OPeNDAP Dataset Access Form

Action: [Get ASCII](#) [Get Binary](#) [Show Help](#)

Data URL: http://www.meteo.unican.es/tds5/dodsC/system4/System4_Seasonal_15Members.ncml

Global Attributes:

```
Originating_or_generating_Center: European Centre for Medium Range Weather Forecasts (ECMWF) (RSMC)
Originating_or_generating_Subcenter: 0
Conventions: CF-1.6, _Coordinates, UW-1.0
history: Read using CDM IOSP Grib1Collection
featureType: GRID
```

Variables: **t2m: Array of 32 bit Reals [run = 0..359][member = 0..14][time06 = 0..860][lat = 0..240][lon = 0..479]**

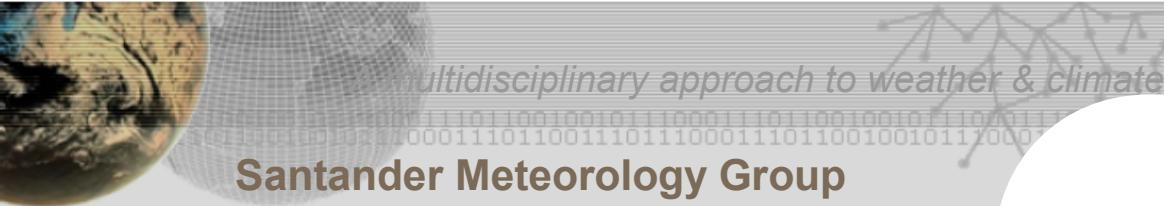
run: member: time06: lat: lon:

```
long_name: 2 metre temperature @ Ground or water surface
units: K
missing_value: NaN
Grib_Variable_Id: VAR_98-0-128-167_L1
Grib1_Center: 98
Grib1_Subcenter: 0
```

mx2t24: Array of 32 bit Reals [run = 0..359][member = 0..14][time24 = 0..215][lat = 0..240][lon = 0..479]

run: member: time24: lat: lon:

```
long_name: Maximum temperature at 2 metres since last 24 hours @
Ground or water surface
units: K
missing_value: NaN
Grib_Variable_Id: VAR_98-0-128-51_L1
Grib1_Center: 98
```



OPeNDAP Dataset Access Form

Action: Get ASCII Get Binary Show Help

Data URL: http://www.meteo.unican.es/tds5/dodsC/system4/System4_Seasonal_15Members.ncml

Global Attributes:

Originating_or_generating_Center: Weather Forecasts (ECMWF) (RSMC)
Originating_or_generating_Subcenter: Conventions: CF-1.6, _Coordinates, history: Read using CDM IOSP Grib1 featureType: GRID

Variables: **t2m: Array of 32 bit Reals [run = 0]**

run: member:
long_name: 2 metre temperature @ G
units: K
missing_value: NaN
Grib_Variable_Id: VAR_98-0-128-167
Grib1_Center: 98
Grib1_Subcenter: 0

mx2t24: Array of 32 bit Reals [run = 0]

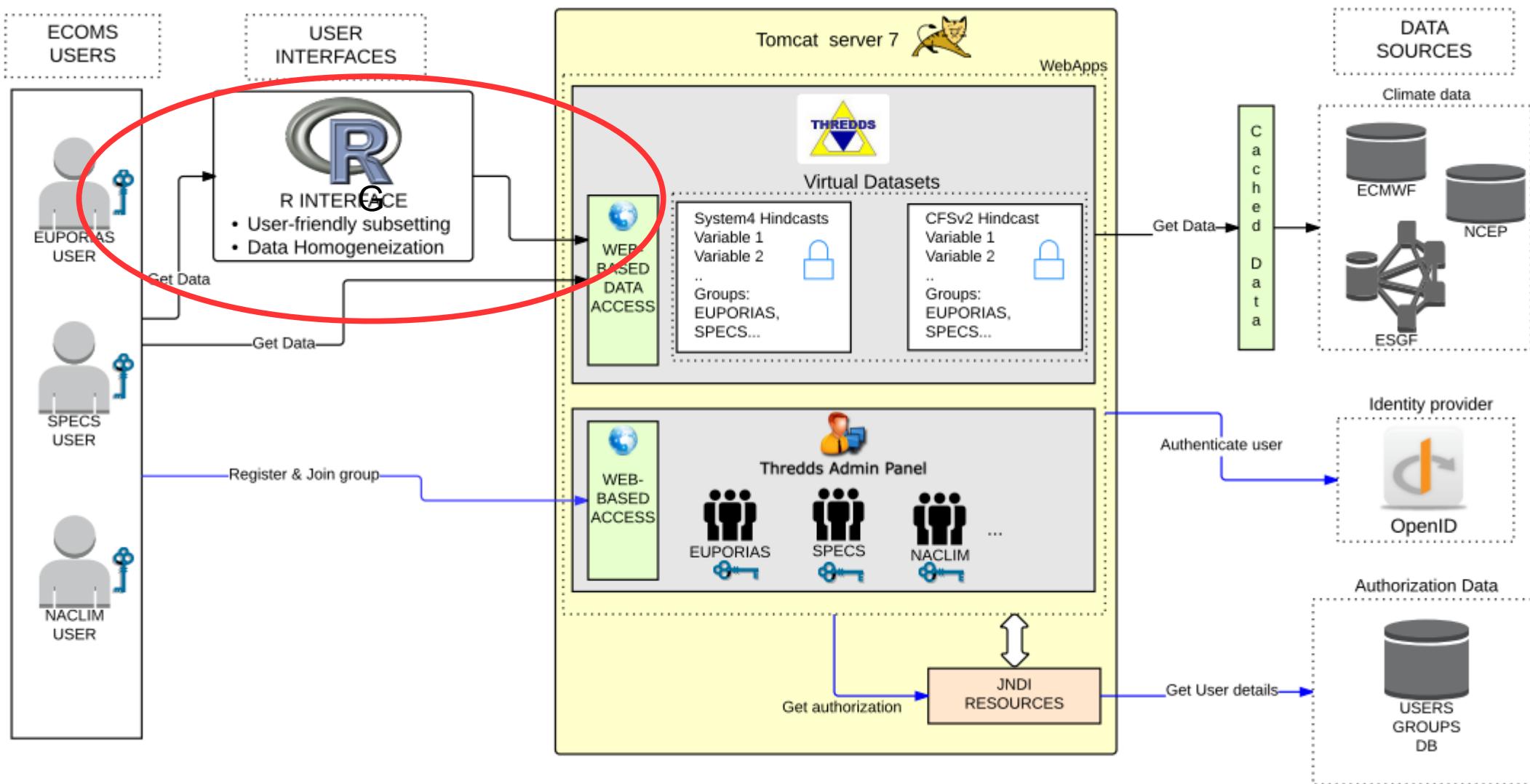
run: 0:1:359 member: 0:1:14
long_name: Maximum temperature at 2
Ground or water surface
units: K
missing_value: NaN
Grib_Variable_Id: VAR_98-0-128-51_L
Grib1_Center: 98

```
Dataset {
    Float32 mx2t24[run = 1][member = 1][time24 = 215][lat = 1]
    [lon = 1];
} system4/System4_Seasonal_15Members.ncml;
-----
mx2t24[1][1][215][1][1]
[0][0][0][0], 272.2452
[0][0][1][0], 271.8321
[0][0][2][0], 272.53897
[0][0][3][0], 273.33136
[0][0][4][0], 273.71527
[0][0][5][0], 273.0291
[0][0][6][0], 273.69446
[0][0][7][0], 273.58426
[0][0][8][0], 274.18262
[0][0][9][0], 273.99835
[0][0][10][0], 273.7032
[0][0][11][0], 273.91562
[0][0][12][0], 274.13705
[0][0][13][0], 273.87527
[0][0][14][0], 274.17377
[0][0][15][0], 274.40704
[0][0][16][0], 274.37482
[0][0][17][0], 274.18506
[0][0][18][0], 273.85782
[0][0][19][0], 273.5874
[0][0][20][0], 274.81393
[0][0][21][0], 273.9504
[0][0][22][0], 273.35272
[0][0][23][0], 273.58716
[0][0][24][0], 274.12602
```

Santander Meteorology Group

A multidisciplinary approach for weather & climate

R Interface



Authentication →

Data load →

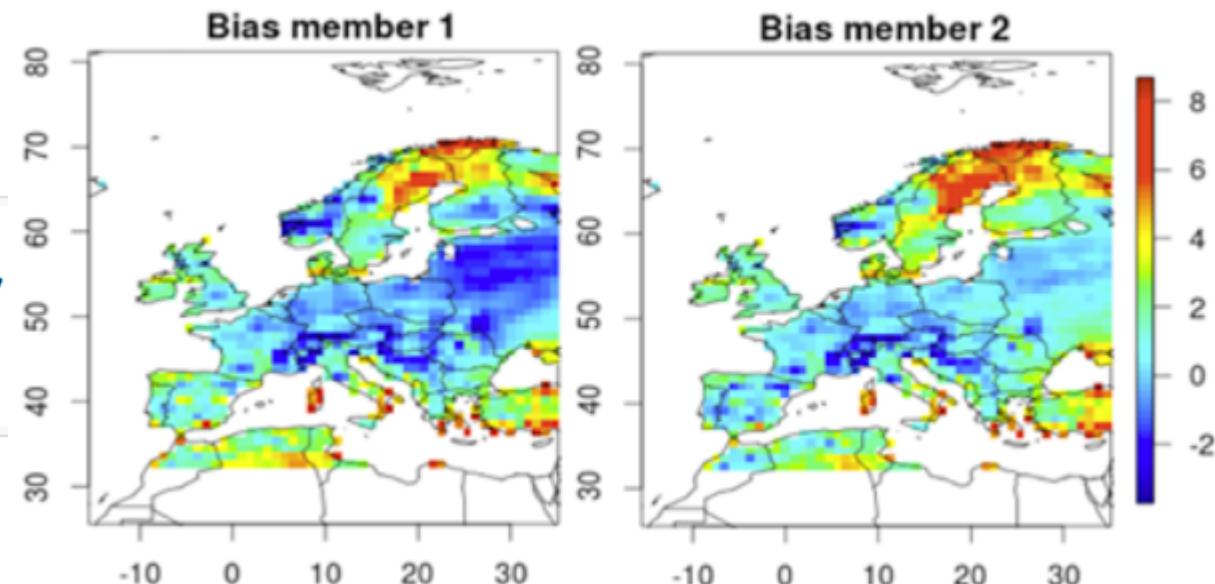
Data processing

- Regridding
- Bias (mock up)
- Plot

```
obsr <- interpGridData(gridData = obs,
                        new.grid = getGrid(prd),
                        method = "bilinear")
bias <- getBias(obsr,prd)
plotMeanField(bias, multi.member = TRUE)
```

```
library(ecomsUDG.Raccess)
loginECOMS_UDG('username',
'password')

obs <- loadECOMS(dataset = "WFDEI",
                  var = "tasmin",
                  season = c(12,1,2))
prd <- loadECOMS(dataset = "System4_seasonal_15",
                  var = "tasmin",
                  season = c(12,1,2),
                  members = 1:2,
                  leadMonth = 1)
```





<http://www.meteo.unican.es/projects/specs/workshop2014>

Santander Meteorology Group
A multidisciplinary approach for weather & climate

Home

- Presentation
- Institutions & location
- Staff
- Teaching activities
- Collaborations
- Contact & travel info

Research

- Research topics
- Projects
- Research networks
- Computing resources
- Publications (stats)
 - Books
 - Papers
 - Proceedings
 - Theses
 - Conferences

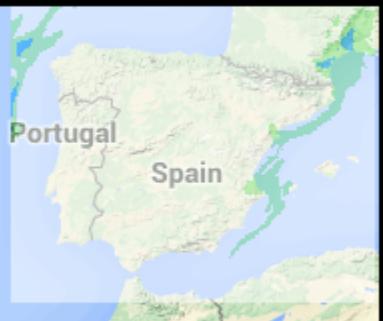
FINAL PROGRAM:

8-9 September 2014 (An introduction to R for climate data analysis)
J Bedia, S Herrera, MD Frias, D SanMartín, M Tuni
[PDF](#) Quick R start: data structures, basic operations, control structures.
PRACTICE: using a climate data object
DEMO: Introduction to R objects [OO-Reference]
PRACTICE: Weather generators

10 September 2014
[09:00-09:15] JM Gutiérrez. [PDF](#) Welcome and presentation
[09:15-10:00] JM Gutiérrez. [PDF](#) Basic concepts on S2D forecasting
[10:00-10:30] M De Felice. [PDF](#) Impact models and data I
[10:30-11:00] K Nicklin. [PDF](#) Impact models and data II
[coffee break]
[11:30-12:00] PA Bretonniere. [PDF](#) SPECS experiments and access
[12:00-12:30] JM Gutiérrez. [PDF](#) ECOMS User Data Gateway
[12:30-13:00] AS Cofiño, J Bedia. Demo: ECOMS-UDG [[DEMO1: Accessing seasonal forecast data using R](#)] [[DEMO2: Validating and visualizing tercile-based probabilistic predictions](#)]
[14:30-17:30] J Bedia, S Herrera. Hands-on training: ECOMS-UDG and R-access package [[PRACTICE1: Bias of System4 hindcast](#)] [[PRACTICE2: Drift of System4 hindcast](#)]

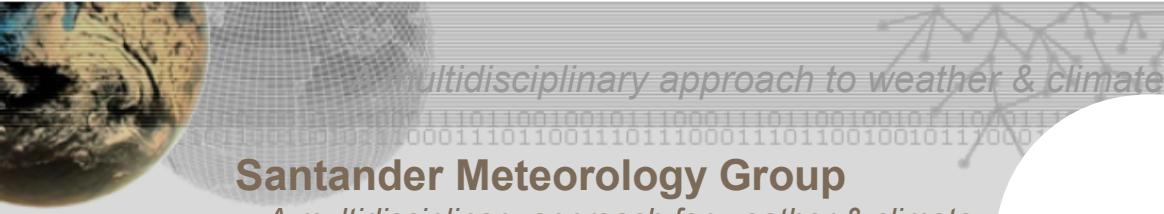
Local weather forecast (physics and statistics) [+ info]
Rainfall (mm) | Max temp. (°C)
Sun (today) Mon Tue Wed


PROMETEO statistical model


WRF dynamical model

For more information [iMeteo](#)

News
2 Sep 2014 [Event]



Some work in progress

- More datasets from different providers.
- **OpenID registration based on ESGF's IdPs**
- **More tools and examples on accessing and processing datasets based on scientific tools/software used by impact community.**
- **Server-side processing capabilities based on WPS.** Job management will be based on meta-scheduling capabilities offered by DRM4G.
- **Remote access on-demand** for ESGF datasets, i.e. SPECS simulations. Improvement of replica/cache strategies.

Thank you!!!

ECOMS User Data Gateway info

<http://meteo.unican.es/ecoms-udg>

ECOMS User Data Gateway access request

<http://meteo.unican.es/tap>

R packages

<https://github.com/SantanderMetGroup/ecomsUDG.Raccess>

<https://github.com/SantanderMetGroup/downscaleR>

Hands-on training workshop material

<http://www.meteo.unican.es/projects/specs/workshop2014>